

Felix Gröner

fgroener@asu.edu

Education

Arizona State University, USA

Expected Fall 2026

Ph.D. Human Systems Engineering

Status: Ph.D. candidate, GPA 4.0

Committee: Erin K. Chiou (Advisor), D. Vaughn Becker, Nancy J. Cooke

Dissertation topic: Engineering AI agents for and evaluating the validity of LLM-driven simulations of synthetic participants for Human Factors research

Ulm University, Germany

June 2024

M.Sc. Cognitive Systems

Thesis: The role of explainable artificial intelligence in trust and adherence to AI recommendations

Kiel University, Germany

December 2022

M.A. International Politics and International Law

Thesis: Necessity of adapting International Humanitarian Law to cyber warfare

Kiel University, Germany

December 2020

B.A. Computer Science and Political Science

Thesis: NATO's strategy to deter cyber attacks

Kiel University, Germany

December 2019

B.Sc. Business Informatics

Thesis: Client's influence on the choice of project management technique in software development

Publications

Gröner, F., Mancenido, M. V., Kim, N., Summers, E., & Chiou, E. K. (in preparation).

Increasing transparency of LLM systems does not always improve people's verification behavior and performance: Results from an empirical study of AI-assisted intelligence analysis reporting.

Gröner, F., & Chiou, E.K. (under review). *Engineering approaches to improve human-LLM interaction: A scoping literature review.*

Gröner, F., & Chiou, E.K. (under review). *Caveats of replacing human participants with LLM-powered simulated users: The case of a writing style preference inference simulation.*

Gröner, F., & Becker, D.V. (in preparation). *Subjective and objective metrics to explain perceptions of original and AI-generated abstract artworks.*

Gröner, F., Mancenido, M. V., & Chiou, E. K. (in preparation). Considering expertise as a factor in the design of synthetic test environments for human-AI research. In E. K. Chiou, D. S. Lange, J. H. Wong, & J. Marble (Eds.), *Advancements in Human-Agent Teaming Research Infrastructure: Testbeds, metrics, and concepts*. CRC Press.

Gröner, F., & Chiou, E. K. (2024). Investigating the impact of user interface designs on expectations about large language models' capabilities. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*.
<https://doi.org/10.1177/10711813241260399>

Kraus, J. M., Merger, J., **Gröner, F.**, & Pätz, J. (2023). 'Sorry' says the robot: The tendency to anthropomorphize and technology affinity affect trust in repair strategies after error. *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (pp. 436-441)*. <https://doi.org/10.1145/3568294.3580122>

Presentations

Chiou, E. K., **Gröner, F.**, & Mancenido, M. V. (2025). Increasing transparency of LLM systems does not always improve people's verification behavior and performance: Results from an empirical study of an AI-assisted intelligence reporting task. Abstract submitted to the *69th HFES International Annual Meeting: ASPIRE*, 13-17 October, Chicago, IL.

Gröner, F., Mancenido, M. V., & Chiou, E. K. (2024). Multi-source AI Scorecard Table (MAST) for evaluating generative AI in worker-automation team tasks. Project update presented at *Center for Accelerating Operational Efficiency (CAOE) Annual Meeting*, 8-9 April, Washington, DC.

Unpublished Projects

Personal project

2025

Semantic leakage is an issue when experimenting with synthetic datasets

When generating synthetic datasets, prompting context can have undesired side-effects which threaten the validity of findings. Mechanistic interpretability methods are used to identify the relevant attention heads. Ablating them did not reduce the issue.

Class project 2025

People and LLMs learning each other's competencies

How well can people and LLMs assess their own and each other's likelihood of knowing the answer to the next question, given their performance history? Sensitivity to prompt design heavily influenced the initial findings.

Master's thesis 2023

The role of explainable artificial intelligence (XAI) in trust and adherence to AI recommendations

Showing people either a global (linear surrogate) or local (LIME) explanation for an AI system's recommendation did not have as large of an effect as the participants' prior self-perceived expertise.

Professional Experience

Center for Accelerating Operational Efficiency 2023-2025

Graduate Research Assistant

Responsibilities: writing grant application and research proposals, authoring papers

Projects:

Graphical user interfaces for directing expectations or LLMs (GUIDE-LLM)

- Development of a custom experimentation platform for human-LLM interaction research. This online platform allows researchers to control the UI/UX of and record behavioral metrics in experiments of human-LLM interaction.
- Composed a set of tasks following the MAST framework to catalog misconception about LLM capabilities.

Multisource AI scorecard table (MAST) for evaluating generative AI in worker-automation team tasks

- Development of a custom experimentation platform for human-LLM interaction research. This online platform allows researchers to control the UI/UX of and record behavioral metrics in experiments of human-LLM interaction.
- Composed a set of tasks following the MAST framework to catalog misconception about LLM capabilities.

The Polytechnic School, Arizona State University 2023

Graduate Teaching Assistant

Responsibilities: Teaching classes and recording videos for IFT101, Raspberry Pi

Human Factors research group, Ulm University 2022-2023

Undergraduate Research Assistant

Responsibilities: Designing and conducting research studies, project management